Analysis of Plasminogen Genetic Variants in Multiple Sclerosis Patients

A Dessa Sadovnick, PhD[*,†], Anthony L Traboulsee, MD[†], Cecily Q Bernales, BSc[*], Jay P Ross, BSc[*], Amanda L Forwell, BSc[*], Irene M Yee, MSc[*], Lena Guillot-Noel, MSc[‡], Bertrand Fontaine, MD, PhD[‡], Isabelle Cournu-Rebeix, PhD[‡], Antonio Alcina, PhD[§], Maria Fedetz, PhD[§,**], Guillermo Izquierdo, MD, PhD[**], Fuencisla Matesanz, PhD[§], Kelly Hilven, MSc[††], Bénédicte Dubois, MD, PhD[††,‡‡], An Goris, PhD[††], Ianire Astobiza, PhD[§§,***], Iraide Alloza, PhD[§§,***,†††], Alfredo Antigüedad, MD[‡‡‡], Koen Vandenbroeck, PhD[§§,***,†††], Denis A Akkad[§§§], Orhan Aktas[****], Paul Blaschke[††††], Mathias Buttmann[‡‡‡‡], Andrew Chan[§§§§], Joerg T Epplen[§§§], Lisa-Ann Gerdes[*****], Antje Kroner[‡‡‡‡,†††††], Christian Kubisch[‡‡‡‡‡], Tania Kümpfel[*****], Peter Lohse[§§§§§,******], Peter Rieckmann[‡‡‡‡,††††††], Uwe K Zettl[††††], Frauke Zipp[‡‡‡‡‡‡], Lars Bertram[§§§§§§,*******], Christina M Lill[‡‡‡‡‡‡,§§§§§§], Oscar Fernandez, MD[†††††††], Patricia Urbaneja, MD[†††††††], Laura Leyva, PhD[‡‡‡‡‡‡‡], Jose Carlos Alvarez-Cermeño, MD[§§§§§§§], Rafael Arroyo, MD, PhD[********], Aroa M Garagorri, MSc[††††††††], Angel García-Martínez MSc[††††††††], Luisa M Villar, PhD[§§§§§§§], Elena Urcelay, PhD[††††††††], Sunny Malhotra, MSc[‡‡‡‡‡‡‡], Xavier Montalban, MD[‡‡‡‡‡‡‡], Manuel Comabella, MD[‡‡‡‡‡‡‡], Thomas Berger, MD[§§§§§§§§], Franz Fazekas, MD[*********], Markus Reindl, MD[§§§§§§§§], Mascha C Schmied, MD[†††††††††], Alexander Zimprich, MD[†††††††††], Carles Vilariño-Güell, PhD[*]

[*] Medical Genetics, University of British Columbia, Vancouver, British Columbia, V6T 1Z3, Canada

[†] Division of Neurology, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, V6T 1Z3, Canada

[‡] Inserm U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle épinière, ICM, Paris, France

[§] Cell Biology and Immunology, Instituto de Parasitología y Biomedicina López Neyra (IPBLN), CSIC, 18100 Granada, Spain

[**] Unidad de Esclerosis Múltiple, Hospital Universitario Virgen Macarena, 41071 Sevilla, Spain

[††] Laboratory for Neuroimmunology, Neurosciences, KU Leuven – University of Leuven, Herestraat 49 bus 1022, Leuven, Belgium

[‡‡] Neurology, University Hospitals Leuven, KU Leuven – University of Leuven, Herestraat 49, Leuven, Belgium

[§§] Neurogenomiks Group, Universidad del País Vasco (UPV/EHU), 48940 Spain

[***] Achucarro Basque Center for Neuroscience, 48170 Spain

[†††] IKERBASQUE, Basque Foundation for Science, 48011 Spain

[‡‡‡] Servicio de Neurología, Hospital Universitario Basurto-Osakidetza, 48940 Bilbao, Spain

[§§§] Human Genetics, Ruhr University, 44801 Bochum, Germany

[****] Neurology, Medical Faculty, Heinrich Heine University, 40225 Düsseldorf, Germany

[††††] Neurology, University of Rostock, 18059 Rostock, Germany

[‡‡‡‡] Neurology, University of Würzburg, 97080 Würzburg, Germany

[§§§§] Neurology, St. Josef-Hospital, Ruhr-University, 44791 Bochum, Germany

[*****] Institute for Clinical Neuroimmunology, Ludwig Maximilian University, 80539 Munich, Germany

[†††††] Centre for Research in Neuroscience at McGill University, Montreal, H3H 2R9 Canada

[‡‡‡‡‡] Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

[§§§§§] Clinical Chemistry, Ludwig Maximilian University, 80539 Munich, Germany

[******] Institute of Laboratory Medicine and Human Genetics, 78224 Singen, Germany

†††††† Neurology, Sozialstiftung Bamberg Hospital, 96049 Bamberg, Germany

‡‡‡‡‡ Neurology, Focus Program Translational Neuroscience, University Medical Center of the Johannes Gutenberg-University Mainz, 55122 Mainz, Germany

§§§§§§ Platform for Genome Analytics, Institutes of Neurogenetics & Integrative and Experimental Genomics, University of Lübeck, 14195 Lübeck, Germany

******* School of Public Health, Medicine, Imperial College London, London SW7 2AZ, UK

††††††† Neurology, UGC Neurociencias Clínicas. IBIMA-Hospital Regional Universitario de Málaga, 29010 Spain

‡‡‡‡‡‡ Research Laboratory, UGC Neurociencias Clínicas. IBIMA-Hospital Regional Universitario de Málaga, 29010 Spain

§§§§§§§ Immunology and Neurology Service, Multiple Sclerosis Unit, Ramón y Cajal University Hospital-IRYCIS, Madrid, Spain

******** Neurology, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos-IdISSC, Madrid, Spain

†††††††† Immunology, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos-IdISSC, Madrid, Spain

‡‡‡‡‡‡‡ Servei de Neurologia-Neuroimmunologia, Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, 08035 Barcelona, Spain

§§§§§§§§ Clinical Neurology, Medical University of Innsbruck, 6020 Austria

********* Neurology, Medical University of Graz, 8010 Austria

††††††††† Neurology, Medical University of Vienna, 1090 Austria

Running Title

Plasminogen in multiple sclerosis

Keywords

Multiple sclerosis, Genetics, Linkage; Association; Plasminogen.

Corresponding author: Carles Vilariño-Güell

Address: 2215 Wesbrook Mall, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

Email: carles@can.ubc.ca

Phone: +1-604-827-1303

Word count: 2947 **Abstract:** 256 **Title characters:** 62

References: 31

Tables: 2 **Figures:** 2

Supplemental Material: 5 Tables and 2 Figures

ABSTRACT

Multiple sclerosis (MS) is a prevalent neurological disease of complex etiology. Here we describe the characterization of a multi-incident MS family which nominated a rare missense variant (p.G420D) in plasminogen (*PLG*) as a putative genetic risk factor for MS. Genotyping of PLG p.G420D (rs139071351) in 2160 MS patients and 886 controls from Canada identified ten additional probands, two sporadic patients and one control with the variant. Segregation analysis in families harboring the rs139071351 variant, identified p.G420D in 26 out of 30 family members diagnosed with MS, 14 unaffected parents and 12 out of 30 family members not diagnosed with disease. Despite considerably reduced penetrance, linkage analysis supports co-segregation of PLG p.G420D and disease. Genotyping of PLG p.G420D in 14446 patients and 8797 controls from Canada, France, Spain, Germany, Belgium and Austria failed to identify significant association with disease (p=0.117), despite an overall higher prevalence in patients (OR=1.32; 95% CI=0.93-1.87). To assess whether additional rare variants have an effect on MS risk, we sequenced *PLG* in 293 probands and genotyped all rare variants in cases and controls. This analysis identified nine rare missense variants, and although three of them were exclusively observed in MS patients segregation analyses do not support pathogenicity. *PLG* is a plausible biological candidate for MS owing to its involvement in immune system response, blood-brain barrier permeability and myelin degradation. Moreover, components of its activation cascade have been shown to present increased activity or expression in MS patients compared to controls; further studies are needed to clarify whether *PLG* is involved in MS susceptibility.

INTRODUCTION

Multiple sclerosis (MS) is a chronic inflammatory demyelinating and neurodegenerative disease of the central nervous system. The genetic contribution to disease susceptibility has been demonstrated in family and twin studies (Ebers et al. 1986; Sadovnick 1993; Fagnani et al. 2015), and the first pathogenic mutation for MS has been recently identified in *NR1H3* (Wang et al. 2016). In addition, a large number genetic risk factors, primarily related to the immune system, have already been identified through association studies (Beecham et al. 2013; Sawcer et al. 2011). However, with the exception of *HLA-DRB1*, all associated variants have a minor effect on overall disease. The identification of genetic components of major effect on disease development is paramount for the generation of physiologically relevant cellular and animal models of human disease, and the generation of treatment strategies that address the underlying biological mechanisms responsible for the onset of MS.

MATERIAL AND METHODS

Participants

A total of 2160 MS patients and 886 unrelated healthy controls from Canada, which includes 1857 multi-incident families, collected through the Canadian Collaborative Project on the Genetic Susceptibility to Multiple Sclerosis (CCPGSMS) were included in this study.(Sadovnick et al. 1998). Five independent European cohorts consisting of 2391 MS patients and 672 healthy controls from France, 4288 patients and 4018 controls from Spain, 3733 patients and 2722 controls from Germany, 1006 patients and 504 controls from Belgium and 925 patients from Austria were used for replication. All patients were diagnosed with MS according to published criteria (Poser et al. 1983; McDonald et al. 2001; Polman et al. 2005), and the demographics for

each cohort are presented in Table 1. The ethical review board at each institution approved the study, and all participants provided written informed consent.

Exome sequencing

We performed exome sequencing in three patients diagnosed with MS (pedigree A; II-1, II-4 and III-1) from a multi-incident family (Figure 1). Exonic regions were enriched using an Ion AmpliSeq exome kit (57.7Mb) and sequenced in an Ion Proton sequencer (Life Technologies, Carlsbad, CA, USA) with a minimum average coverage of 50 reads per base and an average read length of 150 bases. The Ion Torrent Server v4 was used to map reads to NCBI Build 37.1 reference genome using the Torrent Mapping Alignment Program (TMAP) and to identify variants differing from the reference. Sequences with a mapping Phred quality score under 20, fewer than five reads or over 95% strand bias were excluded from further analysis.

Sequencing, genotyping and statistical analysis

Sanger sequencing was used to genotype amplicons containing exome variants of interest and all 19 coding exons and exon-intron boundaries of plasminogen (*PLG*, NM_000301.3) by polymerase chain reaction (PCR) as previously described (Sadovnick et al. 2013). Nine tagging SNPs (tSNPs) spanning a 61 kb region encompassing the *PLG* locus were selected based on HapMap data (version 3 release 27) using Haploview software (Barrett et al. 2005). Selected tSNPs captured over 92% of the polymorphic variation in the region (minor allele frequency (MAF)>5% and r^2>0.8) in Caucasian population standards. Genotyping of variants was performed using a combination of TaqMan probes and Sequenom MassArray iPLEX as previously described (Traboulsee et al. 2014; Nishioka et al. 2010). Genotyping success rate was over 99.4% for all variants and without deviation from Hardy-Weinberg equilibrium expectation (p-value > 0.005). Statistical association was determined using logistic regression analysis

adjusted for age and gender, in addition the combined cohort analysis was adjusted for site. Genotypes were dichotomized as presence versus absence of the minor allele (dominant model). The combined dataset was obtained by pooling samples from all populations. Segregation analysis was quantified using non-parametric and parametric linkage analysis. Non-parametric linkage analysis were performed using SimWalk2 software (version 2.91) and NPL-All statistic (Sobel et al. 2001). Two-point parametric logarithm of odds (LOD) score were obtained with MLINK assuming a dominant model, with a fully penetrant disease and without phenocopies (Ott 1989). All MS patients were treated as affected, non-carrier individuals as healthy and unaffected mutation carriers were treated as having an unknown disease status. The deleterious allele was defined with a 0.0001 frequency and the marker-allele frequency determined empirically from genotyped individuals.

Haplotype analysis

Microsatellite markers spanning the *PLG* locus between D6S1633 and D6S297 were chosen to define the disease-carrying haplotype (Supplementary Table 1). All family members from those families identified with the PLG p.G420D mutation were genotyped. One primer for each pair was labeled with a fluorescent tag and PCR reactions were performed under standard conditions. PCR products were run on an ABI 3730xl (Life Technologies, Carlsbad, CA, USA), and analyzed using GeneMapper 4.0. Marker sizes were normalized to those reported in the CEPH database and manually phased within each family.

RESULTS

To identify genes and variants of major effect on MS susceptibility, we applied exome sequencing analysis to a multi-incident family consisting of 12 individuals over three generations, with DNA available for nine family members, including six diagnosed with MS

(Figure 1A). Exome analysis of II-1, II-4 and III-1 identified 47479, 46545 and 46580 variants, respectively. Of those, 25 missense variants with a MAF below 1% from public and proprietary databases of variants were identified in all three individuals (Supplementary Table 2). Segregation analysis in additional family members identified ten variants shared amongst at least five of the six family members diagnosed with MS for whom DNA was available, and no more than one of the two unaffected blood relatives. Three of these variants were subsequently excluded as they were identified at a frequency over 1% in 366 ethnically matched controls (Supplementary Table 2). The seven remaining variants were genotyped in a multi-ethnic cohort consisting of 2160 MS patients and 886 unrelated healthy controls from Canada. Three variants (TGFBI, p.V608L (ss1467426521); SPINK13, p.C72R (ss1467426567); OR1E1, p.D96Y (ss1467426912)) appear to be private as they were not observed in any of the other samples genotyped in this study and have not been described in public databases of variants (Abecasis et al. 2012; Exome Aggregation Consortium et al. 2015). ARHGAP10, p.T518K (rs375188932), with a reported MAF of $5x10^{-5}$ in the ExAC database, was also not observed in any additional samples. Segregation of these four variants within the exome sequenced family is provided in supplementary figure 1. Of the remainder, SPATA18 p.P286L (rs150116592) was identified in two MS patients, UNC45B p.R776Q (rs34242925) was identified in one patient and one control, and PLG p.G420D (rs139071351) in 12 MS patients and one control.

Segregation analysis for variants identified in *SPATA18* and *UNC45B* did not support co-segregation with disease in additional families and were excluded from further analysis (Supplementary Figure 1). Segregation analysis of PLG p.G420D identified the variant in 26 out of 30 family members diagnosed with MS (87%), 14 parents of MS patients (including eight obligate carriers) not known to suffer from MS, and 12 out of 30 family members not diagnosed

with disease (Figure 1B-M). To quantifiably assess segregation we performed nonparametric and parametric linkage analysis for PLG p.G420D. The more conservative nonparametric score resulted in a LOD score of 1.29, whereas parametric linkage analysis resulted in a maximum LOD score of 5.48 (theta=0.05), despite a penetrance estimate of 50%. Additional support for a role in disease susceptibility is provided by the level of conservation for the glycine residue in mammals, indicating the importance of this amino acid for protein function (Figure 2). Haplotype analysis of PLG p.G420D carriers between D6S1633 and D6S297 did not identify a shared haplotype amongst families (Supplementary Table 1), thus suggesting that PLG p.G420D is a mutational hotspot that has independently arisen in each family rather than being inherited from a common ancestor.

Clinical details were available for 17 PLG p.G420D carriers, five males and 12 females (Supplementary Table 3). The disease course observed in these carriers was predominantly consistent with relapsing-remitting MS or secondary progressive MS with only two patients presenting primary progressive MS. On average, the age at onset of disease was 35.1 years (SD ± 9.1) with a disease duration of 19.9 years (SD ± 10.4). Disease severity was overall relatively moderate, with an average expanded disability status scale (EDSS) score of 3.92 (SD ± 2.9) and a median of 2.75.

Association analysis of PLG p.G420D was performed in Caucasian samples from Canada already genotyped for the identification of additional PLG p.G420D families. This subset consists of 2103 MS patients and 881 controls, and resulted in a marginally significant association with disease risk (p=0.046) and an odds ratio (OR) of 10.19 (Table 1). In order to validate this association we genotyped PLG p.G420D in five independent cohorts from Europe consisting of 12343 MS patients and 7916 healthy controls. Logistic regression analysis

corrected for age and gender identified a similarly marginal association with disease in the French cohort (p=0.049; OR=2.69) whereas no association was observed for any additional cohort (Table 1). Although the combined dataset did not result in a significant association with disease risk (p=0.117), with the exception of Belgium which is the smallest set, all cohorts resulted in OR greater than 1, indicating a higher prevalence of PLG p.G420D in MS patients than controls.

To assess whether common variants in *PLG* lead to an increased susceptibility to develop MS, we identified nine tSNPs spanning the entire *PLG* loci and genotyped them in 2103 MS patients and 881 controls from Canada (Supplementary Table 4). Association analysis failed to identify a significant association between any of the tSNPs and susceptibility to MS ($p>0.05$). Since common variants in *PLG* do not appear to have an effect on MS disease risk, we assessed for the presence of additional rare PLG substitutions in MS patients. To this end we sequenced all *PLG*-coding exons in 293 familial probands from Canada, which identified 11 silent and 11 missense variants (Supplementary Table 5). Of those, nine missense variants with a MAF below 1% in at least two of three publicly available databases (1000G, ExAC or ESP) were genotyped in cases and controls from Canada (Abecasis et al. 2012; Exome Aggregation Consortium et al. 2015; Exome Sequencing Project). This analysis identified six variants (p.K38E, p.R89K, p.R261H, p.R490Q, p.A494V and p.R523W) at similar frequencies in MS patients and controls; whereas p.T200A (rs149145958), p.T500M (rs140970354) and p.A507V (rs372603134) were only identified in eight, two and one MS patient, respectively (Table 2). Despite all three variants being predicted likely damaging to protein function with a phred-scaled CADD score of 29.3, 14.4 and 18.9 for p.T200A, p.T500M and p.A507V, respectively (Kircher et al. 2014); and two of them being evolutionarily conserved (Figure 2), segregation and parametric linkage analysis,

which resulted in negative LOD scores, does not support a role for these variants in disease pathogenicity (Supplementary Figure 2).

DISCUSSION

Exome sequencing analysis in a multi-incident family suffering from MS has nominated PLG p.G420D as a putative new risk factor for MS. Although four private missense variants cannot be conclusively excluded as a potential cause of disease in this kindred, and copy number changes were not evaluated, the identification of PLG p.G420D in twelve additional MS patients and one control from Canada suggests a role for *PLG* in MS susceptibility. Genotyping of additional family members from multi-incident families with PLG p.G420D resulted in positive co-segregation of the variant and disease, albeit with 50% reduced penetrance (Figure 1). Additional support for pathogenicity was sought from a large case-control cohort of MS patients from Europe; and although most populations present a higher prevalence of PLG p.G420D in MS patients than controls, a nominally significant difference was only observed in the French cohort (Table 1). A possible Acadian origin of PLG p.G420D was considered due to the marginal associations in the French and Canadian population; however, the wide geographical distribution of variant carriers from Canada and the lack of a shared ancestral haplotype (Supplementary Table 1) do not support this hypothesis. Association analysis for PLG p.G420D in the entire cohort resulted in a non-significant p-value of 0.117 and an OR of 1.32. Despite the overall lack of association observed, it is possible that carriers of the PLG p.G420D variant have an increased risk of developing MS, as suggested by the OR and initially observed familial segregation pattern. In contrast, common *PLG* tagging variants genotyped in this study were clearly not associated with MS risk in the Canadian population (Supplementary Table 4). This data

corroborates previously described genome wide association studies which did not nominate common variants in *PLG* as a risk factor for MS (Beecham et al. 2013; Sawcer et al. 2011).

Sequencing of *PLG* in MS patients from Canada led to the identification of nine rare missense variants (Table 2). Six of which were subsequently identified at a similar frequency in MS patients and controls, suggesting they are not likely to have an effect on MS risk. Interestingly one of these variants (p.K38E, rs73015965) has been described as the cause of PLG deficiency type I when identified in homozygous or compound heterozygous form (Tefs et al. 2006). Similarly, p.R523W (rs4252129) has been associated with decreased plasma PLG levels (Ma et al. 2014). Severe PLG deficiency type I has been causally linked to ligneous conjunctivitis, a rare chronic inflammatory disease of mainly mucous membranes. Although there is no indication that heterozygous carriers are at an increased risk of developing disease (Tefs et al. 2006), *PLG* dysregulation could lead to an increased susceptibility to inflammatory and autoimmune diseases. In our study, three additional variants (p.T200A, p.T500M and p.A507V) not known to cause hypoplasminogenemia, were observed exclusively in MS patients. Although the allelic frequencies and segregation analysis for rare missense PLG variants do not initially support a role in disease susceptibility, genotyping in additional MS patients is warranted to fully define these preliminary findings. PLG p.T200A seems of particular interest, as it was identified in eight MS patients and no controls (Table 2), it is evolutionary conserved (Figure 2), and a threonine to proline substitution at the same position has been identified in a patient with severe type I PLG deficiency (Tefs et al. 2006).

PLG is a plausible biological candidate for MS susceptibility as it is involved in the inflammatory response, blood-brain barrier (BBB) permeability, neuronal viability, and myelin degradation (Syrovets et al. 2012; Yao and Tsirka 2011; Chen and Strickland 1997; Cuzner and

Opdenakker 1999). PLG has been shown to play a role in the immune response, with plasmin deficiency, the active form of PLG, resulting in a compromised inflammatory response in mouse brain (Hultman et al. 2014). Microglia and astrocytes are the primary mediators of inflammation in the central nervous system, and fibrin has been shown to activate their immune response by stimulating the production of inflammatory mediators including proinflammatory cytokines and reactive oxygen species, as well as act as a chemoattractant for immune cells (Syrovets et al. 2012; Hultman et al. 2014).

Genetic variants in *PLG* may also have an effect on brain inflammation by altering the BBB permeability. Plasmin alters BBB permeability by inducing morphological changes in brain astrocytes and endothelial cells through the reorganization of the actin cytoskeleton and the redistribution of tight junction proteins (Niego and Medcalf 2014; Yao and Tsirka 2011). In addition to its effects on the inflammatory response and BBB permeability, plasmin has also been shown to affect neuronal viability, including sprouting, plasticity and extracellular matrix-related neuronal death (Chen and Strickland 1997; Nakagami et al. 2000; Wu et al. 2000).

Plasmin activates highly active matrix metalloproteinases (MMPs) which are recognized as key proteases in the demyelination process. Synthetic inhibitors of MMPs have been found to ameliorate clinical symptoms and pathological signs in experimental autoimmune encephalomyelitis (EAE) animal models (Cuzner and Opdenakker 1999); and minocycline, which has several immunomodulating activities including the inhibition of MMP-9, has being successfully used in clinical trials as an add-on therapy for MS patients (Metz et al. 2009).

Despite the existence of extended families with a high incidence of MS (Fagnani et al. 2015; Sadovnick 1993), only one rare pathogenic mutations has been reported (Wang et al. 2016). In this study, the implementation of exome sequencing analysis in a multi-incident MS family

nominated PLG p.G420D as a potential susceptibility risk for MS. Additional support was provided by 10 additional multi-incident MS families in which the variant segregates with disease, albeit with reduced penetrance. Disappointingly, genotyping of PLG p.G420D in a large European case-control cohort failed to identify a significant association with MS, thus not supporting a role in disease. Despite this lack of association, dysregulation of the PLG/plasmin activation cascade is a plausible pathomechanism of MS; which in conjunction with the positive segregation of PLG p.G420D in families (Figure 1), the overall higher incidence of PLG p.G420D carriers in European MS patients (Table 1), and the identification of additional rare PLG substitutions in MS patients not observed in controls (Table 2), warrants further genetic and functional characterization of *PLG* in order to elucidate its potential role on MS susceptibility and pathogenesis.

ACKNOWLEDGEMENTS

We are grateful to all individuals who generously participated in this study. We thank Kevin Atkins for data collection and extraction. We also thank Généthon, AFM, ARSEP and CRB-REFGENSEP.

FUNDING

This research was undertaken thanks to funding from the Canada Research Chair [950-228408] and Canada Excellence Research Chair programs [214444], Canadian Institutes of Health Research [MOP-137051], Vancouver Coastal Health Research Institute, the Milan & Maureen Ilich Foundation [11-32095000], and the Vancouver Foundation [ADV14-1597]. Replication

studies received funding from the program “Investissements d’avenir” ANR-10-IAIHU-06. Fondo de Investigación Sanitaria (FIS)-Instituto de Salud Carlos III(ISCIII)-Fondos Europeos de Desarrollo Regional (FEDER), Unión Europea [grant numbers P12/00555, PI13/01527, PI13/01466 and PI13/0879 to FM, AA and GI] and Junta de Andalucía -FEDER [grant number CTS2704 to FM]. BD is a Clinical Investigator of the Research Foundation Flanders (FWO-Vlaanderen). AG and BD are supported by the Research Fund KU Leuven (OT/11/087 and CREA/14/023) and the Research Foundation Flanders (G073415N).

AUTHORS CONTRIBUTIONS

CV-G, ADS, ALT, IC-R, FM, BD, AG, KV, LB, CML, LL, EU, MC and AZ conceived and designed the experiments. CQB, JPR and ALF performed the experiments. CV-G, IC-R, FM, BD, AG, KV, LB, CML, LL, EU, MC and AZ analyzed the data. CV-G, ADS, ALT, IMY, LG-N, BF, IC-R, AnA, MF, GI, FM, KH, BD, AG, IaA, IrA, AlA, KV, DAA, OA, PB, MB, AC, JTE, LG, AK, CK, TK, PL, PR, UKZ, FZ, LB, CML, OF, PU, LL, JCA-C, RA, AMG, AG-M, LMV, EU, SM, XM, MC, TB, FF, MR, MCS and AZ contributed reagents/materials/analysis tools. CV-G, ADS, ALT, IMY, IC-R, AnA, FM, KH, BD, AG, IrA, AlA, KV, DAA, JTE, MB, CK, LB, CML, LL, EU, MC, TB, FF, MR and AZ contributed to the writing of the manuscript.

COMPETING INTERESTS

ALT reports personal fees from Biogen Idec, Chugai, Medimmune, Teva Innovation, and EMD Serono, and grants and personal fees from Genzyme Sanofi and Roche. All other authors report no disclosures.

FIGURE LEGENDS

Figure 1. Simplified pedigrees for families presenting the PLG p.G420D variant. Males are represented by squares and females by circles, the proband is indicated with an arrow head. Patients diagnosed with MS have black filled symbols and carriers of unknown clinical phenotype have grey filled symbols. Heterozygote carriers (M) and wild-type (wt) genotypes are indicated. An asterisk indicates an inferred carrier. Pedigree A was used for exome analysis, and with the exception of pedigree E, which is of Asian descent, all families are of Caucasian ancestry.

Figure 2. PLG variants and cross-species conservation. Protein orthologs were aligned via ClustalO. Amino acid positions for PLG variants are highlighted in black. Protein orthologs with amino acid positions differing from those of the human sequence are indicated in gray. RefSeq accession numbers: *Homo sapiens* NP_000292.1, *Macaca mulatta* NP_001036540.1, *Mus musculus* NP_032903.3, *Rattus norvegicus* NP_445943.1, *Canis lupus familiaris* NP_001273889.1, *Sus scrofa* NP_001038055.1, *Bos taurus* NP_776376.1, *Myotis davidii* ELK34830.1, *Tarsius syrichta* XP_008066085.1, *Gallus gallus* XP_419618.2, *Danio rerio* AAH59801.1.

REFERENCES

Abecasis, G. R., A. Auton, L. D. Brooks, M. A. DePristo, R. M. Durbin *et al.*, 2012 An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56-65.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly, 2005 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263-265.

Beecham, A. H., N. A. Patsopoulos, D. K. Xifara, M. F. Davis, A. Kemppinen *et al.*, 2013 Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 45: 1353-1360.

Chen, Z. L., and S. Strickland, 1997 Neuronal death in the hippocampus is promoted by plasmin-catalyzed degradation of laminin. Cell 91: 917-925.

Cuzner, M. L., and G. Opdenakker, 1999 Plasminogen activators and matrix metalloproteases, mediators of extracellular proteolysis in inflammatory demyelination of the central nervous system. J Neuroimmunol 94: 1-14.

Ebers, G. C., D. E. Bulman, A. D. Sadovnick, D. W. Paty, S. Warren *et al.*, 1986 A population-based study of multiple sclerosis in twins. N Engl J Med 315: 1638-1642.

Exome Aggregation Consortium, M. Lek, K. Karczewski, E. Minikel, and K. Samocha, 2015 Analysis of protein-coding genetic variation in 60,706 humans. BioRxiv: doi: http://dx.doi.org/10.1101/030338.

Exome Sequencing Project, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [2014].

Fagnani, C., M. C. Neale, L. Nistico, M. A. Stazi, V. A. Ricigliano *et al.*, 2015 Twin studies in multiple sclerosis: A meta-estimation of heritability and environmentality. Mult Scler: pii: 1352458514564492.

Hultman, K., M. Cortes-Canteli, A. Bounoutas, A. T. Richards, S. Strickland *et al.*, 2014 Plasmin deficiency leads to fibrin accumulation and a compromised inflammatory response in the mouse brain. J Thromb Haemost 12: 701-712.

Kircher, M., D. M. Witten, P. Jain, B. J. O'Roak, G. M. Cooper *et al.*, 2014 A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46: 310-315.

Ma, Q., A. B. Ozel, S. Ramdas, B. McGee, R. Khoriaty *et al.*, 2014 Genetic variants in PLG, LPA, and SIGLEC 14 as well as smoking contribute to plasma plasminogen levels. Blood 124: 3155-3164.

McDonald, W. I., A. Compston, G. Edan, D. Goodkin, H. P. Hartung *et al.*, 2001 Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol 50: 121-127.

Metz, L. M., D. Li, A. Traboulsee, M. L. Myles, P. Duquette *et al.*, 2009 Glatiramer acetate in combination with minocycline in patients with relapsing--remitting multiple sclerosis: results of a Canadian, multicenter, double-blind, placebo-controlled trial. Mult Scler 15: 1183-1194.

Nakagami, Y., K. Abe, N. Nishiyama, and N. Matsuki, 2000 Laminin degradation by plasmin regulates long-term potentiation. J Neurosci 20: 2003-2010.

Niego, B., and R. L. Medcalf, 2014 Plasmin-dependent modulation of the blood-brain barrier: a major consideration during tPA-induced thrombolysis? J Cereb Blood Flow Metab 34: 1283-1296.

Nishioka, K., C. Wider, C. Vilarino-Guell, A. I. Soto-Ortolaza, S. J. Lincoln *et al.*, 2010 Association of alpha-, beta-, and gamma-Synuclein with diffuse lewy body disease. Arch Neurol 67: 970-975.

Ott, J., 1989 Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci U S A 86: 4175-4178.

Polman, C. H., S. C. Reingold, G. Edan, M. Filippi, H. P. Hartung *et al.*, 2005 Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol 58: 840-846.

Poser, C. M., D. W. Paty, L. Scheinberg, W. I. McDonald, F. A. Davis *et al.*, 1983 New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 13: 227-231.

Sadovnick, A. D., 1993 Familial recurrence risks and inheritance of multiple sclerosis. Curr Opin Neurol Neurosurg 6: 189-194.

Sadovnick, A. D., N. J. Risch, and G. C. Ebers, 1998 Canadian collaborative project on genetic susceptibility to MS, phase 2: rationale and method. Canadian Collaborative Study Group. Can J Neurol Sci 25: 216-221.

Sadovnick, A. D., A. L. Traboulsee, J. D. Lee, J. P. Ross, C. Q. Bernales *et al.*, 2013 Colony stimulation factor 1 receptor (CSF1R) is not a common cause of multiple sclerosis. Eur J Neurol 20: e115-116.

Sawcer, S., G. Hellenthal, M. Pirinen, C. C. Spencer, N. A. Patsopoulos *et al.*, 2011 Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476: 214-219.

Sobel, E., H. Sengul, and D. E. Weeks, 2001 Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. Hum Hered 52: 121-131.

Syrovets, T., O. Lunov, and T. Simmet, 2012 Plasmin as a proinflammatory cell activator. J Leukoc Biol 92: 509-519.

Tefs, K., M. Gueorguieva, J. Klammt, C. M. Allen, D. Aktas *et al.*, 2006 Molecular and clinical spectrum of type I plasminogen deficiency: A series of 50 patients. Blood 108: 3021-3026.

Traboulsee, A. L., C. Q. Bernales, J. P. Ross, J. D. Lee, A. D. Sadovnick *et al.*, 2014 Genetic variants in IL2RA and IL7R affect multiple sclerosis disease risk and progression. Neurogenetics 15: 165-169.

Wang, Z., A. D. Sadovnick, A. L. Traboulsee, J. P. Ross, C. Q. Bernales *et al.*, 2016 Nuclear receptor NR1H3 in familial multiple sclerosis. Neuron: DOI: 10.1016/j.neuron.2016.1004.1039.

Wu, Y. P., C. J. Siao, W. Lu, T. C. Sung, M. A. Frohman *et al.*, 2000 The tissue plasminogen activator (tPA)/plasmin extracellular proteolytic system regulates seizure-induced hippocampal mossy fiber outgrowth through a proteoglycan substrate. J Cell Biol 148: 1295-1304.

Yao, Y., and S. E. Tsirka, 2011 Truncation of monocyte chemoattractant protein 1 by plasmin promotes blood-brain barrier disruption. J Cell Sci 124: 1486-1495.

Table 1. Logistic regression analysis for PLG p.G420D (rs139071351) and risk of MS. F, female; M, male; OR, odds ratio; CI, confidence interval; NA, not applicable.

	Group	Gender F:M	Age (mean ± SD)	Age at onset (mean ± SD)	Genotypes (GA/GG)	p-value	OR (95% CI)
Canada	Controls	1:1.04	67.1 ± 9.8		1/880	0.046	10.19 (1.04-267.89)
	MS patients	1:0.37	46.7 ± 11.7	31.0 ± 9.7	12/2091		
France	Controls	1:0.64	39.3 ± 13.1		4/668	0.049	2.69 (1.00-9.37)
	MS patients	1:0.43	49.1 ± 11.4	30.5 ± 9.7	32/2359		
Spain	Controls	1:0.68	42.8 ± 12.8		34/3984	0.475	1.20 (0.73-1.96)
	MS patients	1:0.53	44.5 ± 11.5	30.9 ± 9.8	42/4246		
Germany	Controls	1:0.68	41.3 ± 16.8		11/2711	0.476	1.31 (0.63-2.84)
	MS patients	1:0.41	40.5 ± 11.3	30.8 ± 10.3	21/3712		
Belgium	Controls	1:0.89	56.2 ± 14.7		5/499	0.747	0.81 (0.23-3.04)
	MS patients	1:0.52	48.3 ± 13.1	33.3 ± 10.9	6/1000		
Austria	MS patients	1:0.43	49.2 ± 12.1	28.7 ± 9.1	7/918	NA	NA
Combined	Controls	1:0.72	44.3 ± 15.9		55/8742	0.117	1.32 (0.93-1.87)
	MS patients	1:0.45	45.1 ± 12.1	30.9 ± 9.9	120/14326		

Table 2. Case-control frequency for rare missense PLG variants identified in MS patients.

dbSNP ID[a]	Chromosome and position	Nucleotide change	Protein change	Minor allele frequency		
				ExAC[b]	Controls (n)	MS (n)
rs73015965	6:161127501	A/G	p.K38E	0.003	0.006 (10)	0.007 (28)
rs143079629	6:161128812	G/A	p.R89K	0.007	0.010 (16)	0.010 (44)
rs149145958	6:161135876	A/G	p.T200A	0.001	0	0.002 (8)
rs4252187	6:161137790	G/A	p.R261H	0.003	0.007 (12)	0.005 (24)
rs140537724	6:161152807	G/A	p.R490Q	0.001	0.002 (3)	0.002 (9)
rs4252128	6:161152819	C/A	p.A494V	0.008	0.005 (8)	0.005 (20)
rs140970354	6:161152837	C/T	p.T500M	0.0002	0	0.0005 (2)
rs372603134	6:161152858	C/T	p.A507V	0.0001	0	0.0002 (1)
rs4252129	6:161152905	C/T	p.R523W	0.007	0.012 (19)	0.013 (56)

[a]dbSNP Build 138, [b]The Exome Aggregation Consortium (ExAC) database

Figure 1

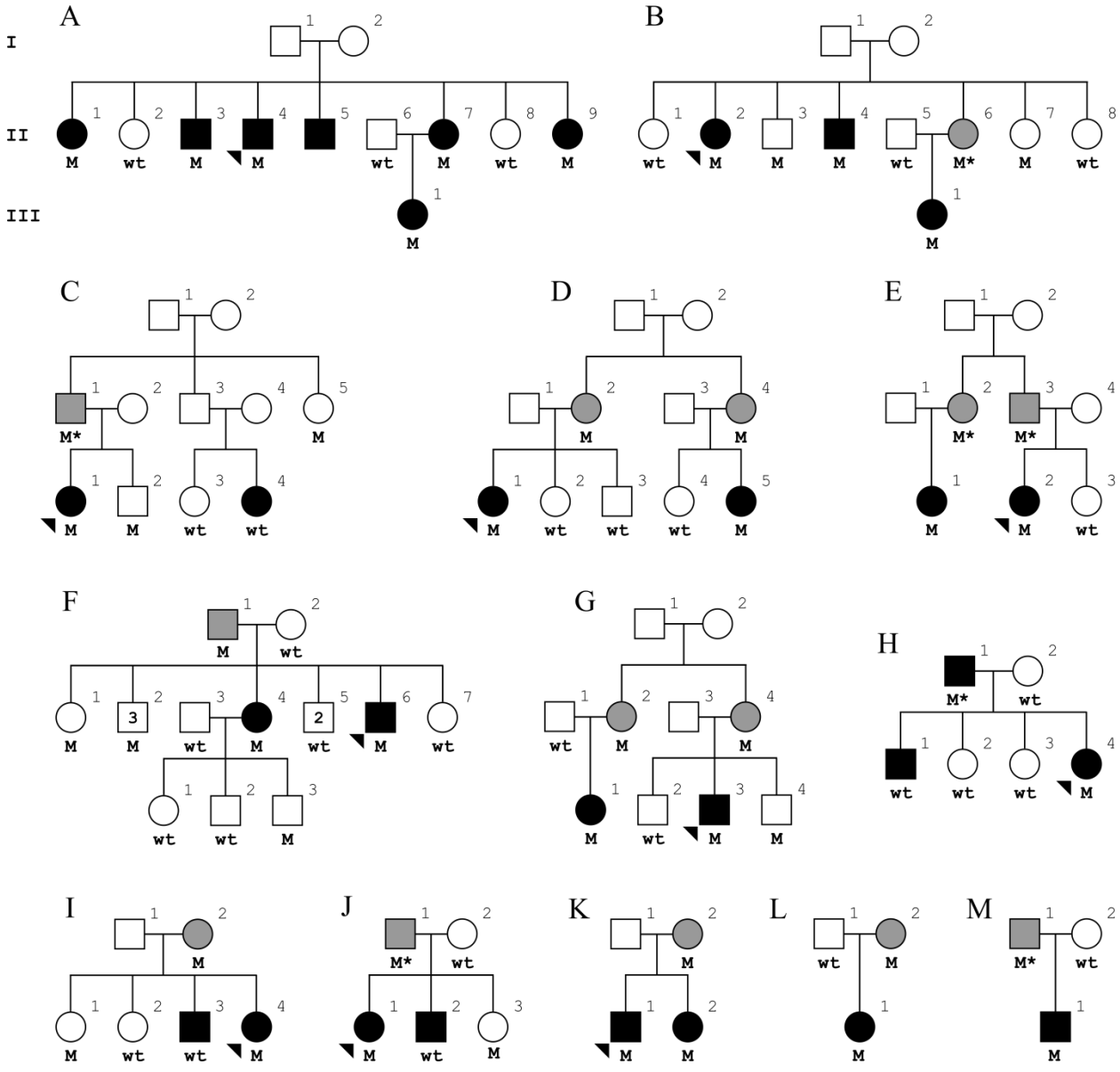
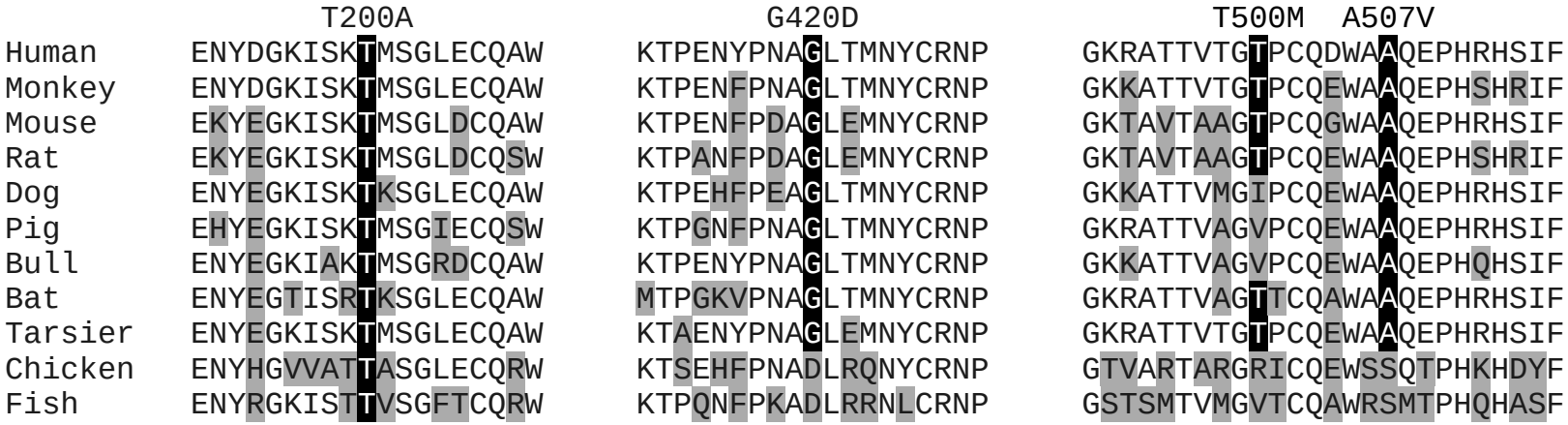

A
I
II
III
M wt M M wt M wt M
M
B
wt M M M wt M* M wt
M
C
M* M
M M wt wt
D
M M
M wt wt wt M
E
M* M*
M M wt
F
M wt
M M 3 wt M 2 M wt
wt wt M
G
wt M M
M wt M M
H
M* wt
wt wt wt M
I
M
M wt wt M
J
M* wt
M wt M
K
M
M M
L
wt M
M
M
M* wt
M

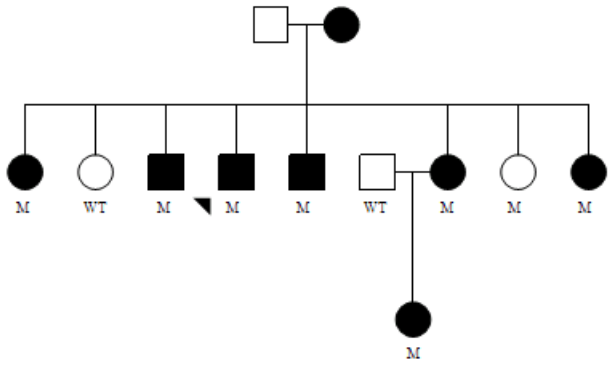
Figure 2.


T200A
G420D
T500M
A507V
Human ENYDGKISKTMSGLECQAW KTPENYPNAGLTMNYCRNP GKRATTVTGTPCQDWAAQEPHRHSIF
Monkey ENYDGKISKTMSGLECQAW KTPENFPNAGLTMNYCRNP GKKATTVTGTPCQEWAAQEPHSHRIF
Mouse EKYEGKISKTMSGLDCQAW KTPENFPDAGLEMNYCRNP GKTAVTAAGTPCQGWAAQEPHRHSIF
Rat EKYEGKISKTMSGLDCQSW KTPANFPDAGLEMNYCRNP GKTAVTAAGTPCQEWAAQEPHSHRIF
Dog ENYEGKISKTKSGLECQAW KTPEHFPEAGLTMNYCRNP GKKATTVMGIPCQEWAAQEPHRHSIF
Pig EHYEGKISKTMSGIECQSW KTPGNFPNAGLTMNYCRNP GKRATTVAGVPCQEWAAQEPHRHSIF
Bull ENYEGKIAKTMSGRDCQAW KTPENYPNAGLTMNYCRNP GKKATTVAGVPCQEWAAQEPHQHSIF
Bat ENYEGTISRTKSGLECQAW MTPGKVPNAGLTMNYCRNP GKRATTVAGTTCQAWAAQEPHRHSIF
Tarsier ENYEGKISKTMSGLECQAW KTAENYPNAGLEMNYCRNP GKRATTVTGTPCQEWAAQEPHRHSIF
Chicken ENYHGVVATTASGLECQRW KTSEHFPNADLRQNYCRNP GTVARTARGRICQEWSSQTPHKHDYF
Fish ENYRGKISTTVSGFTCQRW KTPQNFPKADLRRNLCRNP GSTSMTVMGVTCQAWRSMTPHQHASF

Supplementary Figure 1. Segregation analysis of exome variants. Males are represented by squares and females by circles, with the proband indicated with an arrow head. Patients diagnosed with MS have black filled symbols. Heterozygote carriers (M) and wild-type (WT) genotypes are indicated.

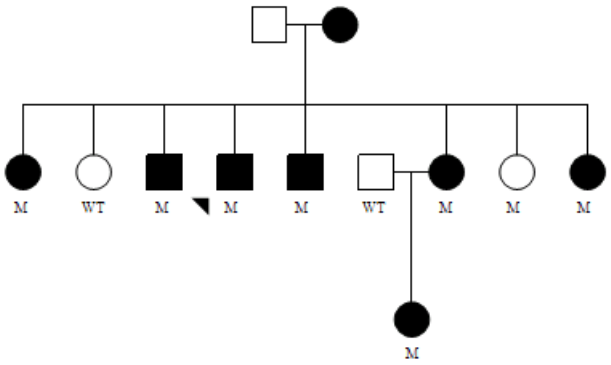
ARHGAP10 p.T518K (rs375188932)

Original family (Fig1 A)



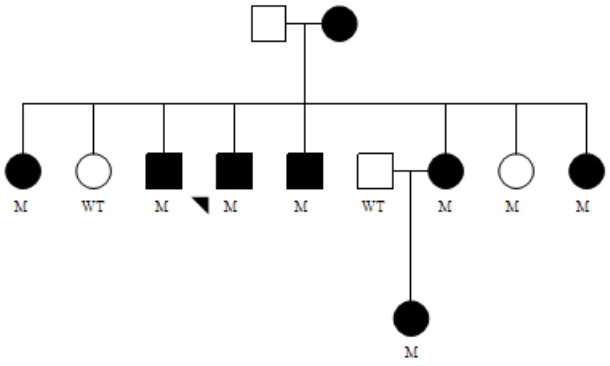
TGFBI p.V608L (ss1467426521)

Original family (Fig1 A)



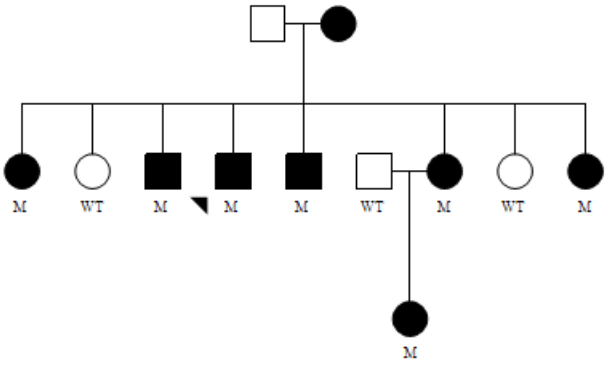
SPINK13 p.C72R (ss1467426567)

Original family (Fig1 A)



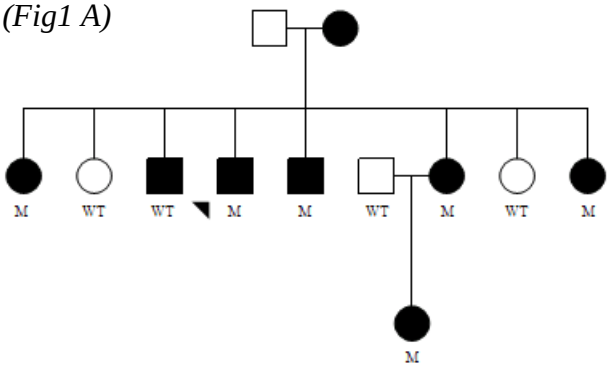
OR1E1 p.D96Y (ss1467426912)

Original family (Fig1 A)

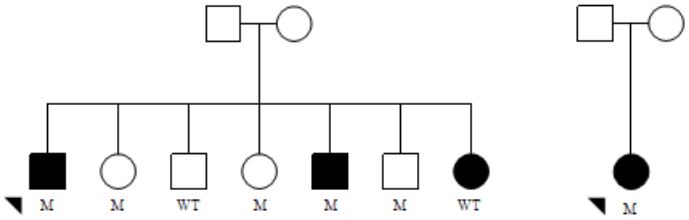


SPATA18 p.P286L (rs150116592)

Original family (Fig1 A)

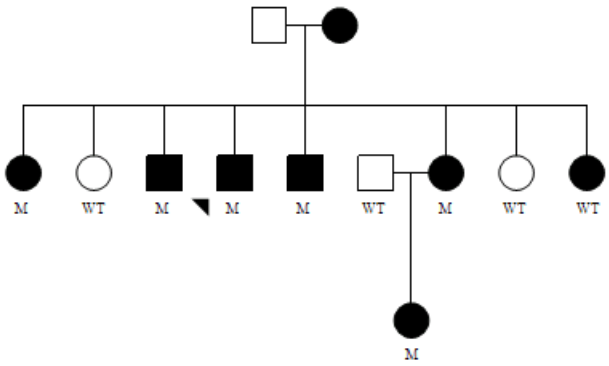


Additional families

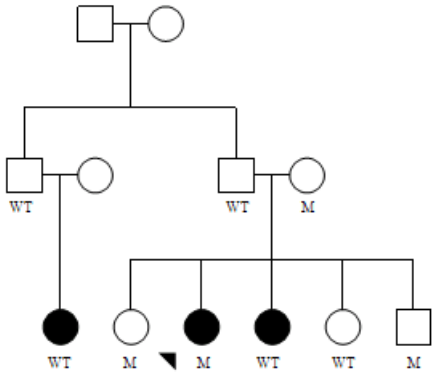


UNC45B p.R776Q (rs34242925)

Original family (Fig1 A)

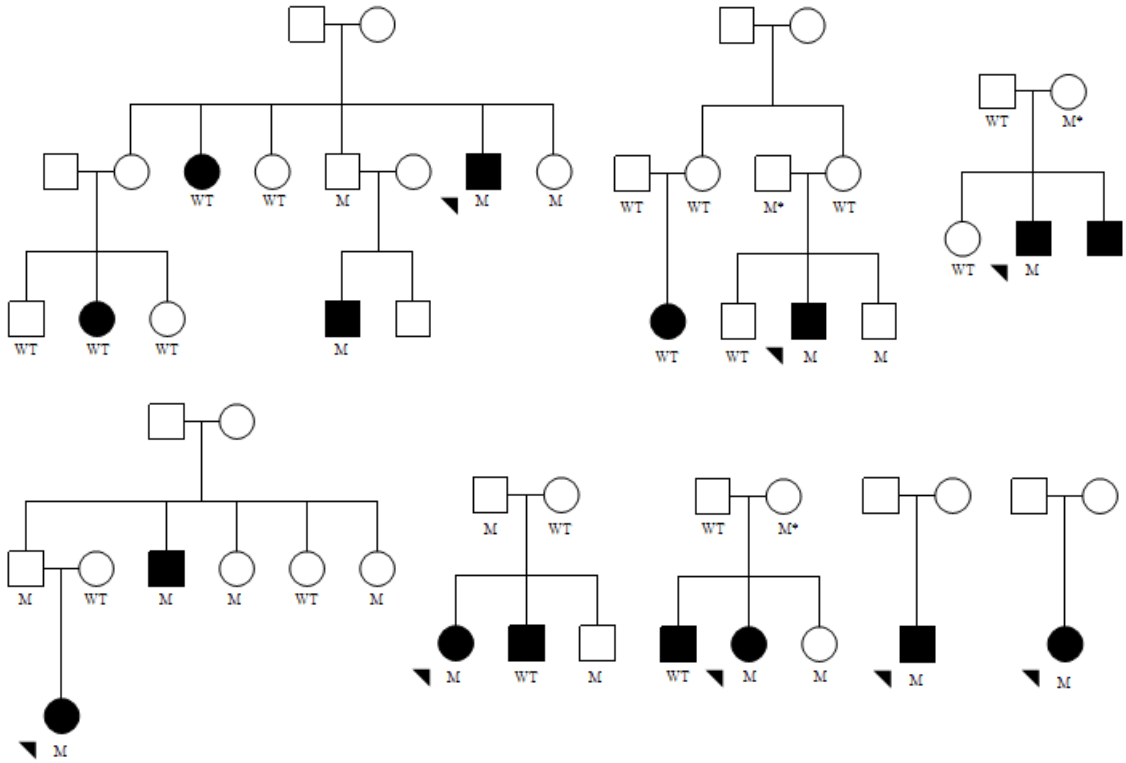


Additional family

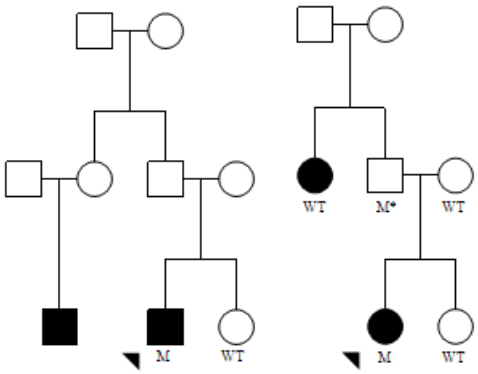


Supplementary Figure 2. Segregation analysis of PLG variants not observed in controls. Males are represented by squares and females by circles, with the proband indicated with an arrow head. Patients diagnosed with MS have black filled symbols. Heterozygote carriers (M) and wild-type (wt) genotypes are indicated. Inferred carriers are indicated with an asterisk.

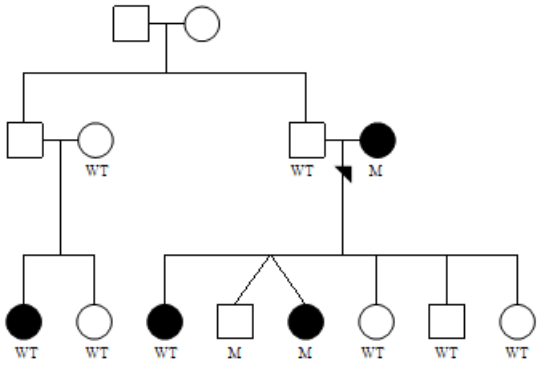
PLG p. T200A (rs149145958)



PLG p.T500M (rs140970354)



PLG p.A507V (rs372603134)



Supplementary Table 1. Chromosome 6q25.3-27 haplotypes carrying PLG p.G420D. Microsatellite allele sizes are given in base pairs consistent with Centre d'Etude du Polymorphisme Humain (CEPH) standards. For markers with an unknown phase both alleles are given. Markers are shown with their physical locations according to NCBI Build 37.1.

Marker	Position	Family ID												
		A	B	C	D	E	F	G	H	J	I	K	L	M
D6S1633	157,016,443	129	125	121	129	119	131	117	121	117	121	121/129	127	121/129
D6S415	157,794,580	263	275	275	265	263	261	261	265	261	265	267	261	263/275
D6S1655	158,300,596	149	149	149/163	165	151	155	149	155	153	155	149	155	149
D6S437	158,683,884	135	153	131	131	161	163	131	147	131	135	131/153	131	131/159
D6S1581	160,196,757	219	229	219	219	219	219	219	219	227	219	219/227	219	215
rs139071351	161,152,085	A	A	A	A	A	A	A	A	A	A	A	A	A
D6S305	162,115,166	218	204	226	224	218	204	218	218	232	204	226	204	226
D6S1599	162,759,584	151	147	131	131	131	133	133	133	143	133	133	133	155
D6S1277	164,217,713	290	296	294/296	296	296	298	302	298	298	298	298	298	298
D6S1719	165,989,655	180	184	182	178	178	180	174	174	182	178	178	178	182
D6S297	167,156,674	222	212	222	214	212	222	212	212	212	222	222	212	212

Supplementary Table 2. Novel and rare (MAF<0.01) missense variants shared by family A members II-1, II-4 and III-1. Chromosomal positions are provided in reference to NCBI Build 37.1.

Chr	Position	Gene	nucleotide change	Transcript	AA change	dbSNP rs/ss number	Selection/exclusion criteria
1	865,665	SAMD11	G/A	NM_152486	R68Q	rs145442390	Does not segregate with disease
1	1,961,467	GABRD	G/A	NM_000815	G369S	rs199865640	Does not segregate with disease
1	226,573,339	PARP1	G/A	NM_001618	L293F	rs149619679	Does not segregate with disease
1	228,505,276	OBSCN	G/T	NM_001098623	R4558L	rs199865640	Does not segregate with disease
1	230,921,738	CAPN9	G/A	NM_016452	G472E	ss1467426444	Does not segregate with disease
4	52,943,043	SPATA18	C/T	NM_145263	P286L	rs150116592	MS frequency = 0.001 Control frequency = 0
4	148,886,277	ARHGAP10	C/A	NM_024605	T518K	rs375188932	Private variant
4	155,254,540	DCHS2	G/C	NM_001142552	H940Q	rs79215995	Control frequency = 0.01
5	40,852,774	CARD6	T/C	NM_032587	L447P	rs143022216	Does not segregate with disease
5	57,790,693	GAPT	A/T	NM_152687	E110D	rs147191680	Does not segregate with disease
5	94,927,243	ARSK	C/T	NM_198150	P337L	rs149766065	Does not segregate with disease
5	98,224,802	CHD1	T/A	NM_001270	Y774F	rs144567251	Does not segregate with disease

5	123,980,164	ZNF608	T/C	NM_020747	K1299R	rs113873110	Control frequency = 0.03
5	135,396,541	TGFBI	G/C	NM_000358	V608L	ss1467426521	Private variant
5	147,661,772	SPINK13	T/C	NM_001040129	C72R	ss1467426567	Private variant
5	176,798,224	RGS14	G/A	NM_006480	R438K	ss1467426609	Does not segregate with disease
6	161,152,085	PLG	G/A	NM_000301	G420D	rs139071351	MS frequency = 0.005 Control frequency = 0.001
10	99,240,767	MMS19	C/T	NM_022362	R65Q	ss1467426865	Does not segregate with disease
15	42,185,151	SPTBN5	G/A	NM_016642	R74C	rs62002144	Does not segregate with disease
15	44,865,000	SPG11	T/C	NM_001160227	N1962S	rs140824939	Does not segregate with disease
16	66,804,108	CCDC79	G/T	NM_001136505	S459R	rs189708354	Does not segregate with disease
17	3,301,419	OR1E1	C/A	NM_003553	D96Y	ss1467426912	Private variant
17	33,507,649	UNC45B	G/A	NM_001033576	R776Q	rs34242925	MS frequency = 0.0005 Control frequency = 0.001
20	68,396	DEFB125	C/T	NM_153325	R16W	rs138777928	Control frequency = 0.01
22	46,780,446	CELSR1	C/T	NM_014246	E2293K	rs140996267	Does not segregate with disease

Supplementary Table 3. Clinical features for PLG p.G420D carriers. RR, Relapsing-remitting MS; PP, primary progressive MS; SP, secondary progressive MS; EDSS, expanded disability status scale; NA, not available.

Pedigree	Individual	Gender	Disease course	Age	Age at onset	Disease duration	EDSS score
A	II-1	F	RR	69	41	28	-
A	II-3	M	PP	65	26	39	-
A	II-4	M	RR	63	44	19	2
A	II-7	F	RR	58	20	38	-
A	II-9	F	RR	52	23	29	-
B	II-2	F	RR	59	35	24	3.5
C	III-1	F	RR	52	37	15	3
D	III-1	F	RR	55	36	19	2
E	III-2	F	SP	53	26	27	8
F	II-6	M	RR	66	54	12	6.5
G	III-1	F	PP	53	45	8	-
G	III-3	M	RR	41	27	14	1
H	II-4	F	RR	50	44	6	1
I	II-4	F	SP	43	31	12	7.5
J	II-1	F	RR	59	37	22	8
L	II-1	F	RR	43	40	3	2.5
M	II-1	M	-	53	30	23	2

Supplementary Table 4. Logistic regression analysis for PLG-tagging SNPs. P-values were corrected for age and gender. Odds ratios (OR) and 95% confidence intervals (CI) are provided on the minor allele.

dbSNP	Genotypes	Controls n (%)			Multiple sclerosis n (%)			p-value	OR (95% CI)
rs9458005	AA/AG/GG	476 (0.58)	297 (0.36)	51 (0.06)	1193 (0.60)	705 (0.35)	98 (0.05)	0.618	0.92 (0.78-1.09)
rs2144723	CC/CT/TT	231 (0.28)	409 (0.50)	182 (0.22)	557 (0.28)	990 (0.50)	448 (0.22)	0.274	1.01 (0.84-1.21)
rs1830519	AA/AG/GG	438 (0.53)	323 (0.39)	61 (0.07)	1120 (0.56)	742 (0.37)	135 (0.07)	0.154	0.89 (0.76-1.05)
rs783147	GG/GA/AA	244 (0.30)	394 (0.48)	184 (0.22)	599 (0.30)	1031 (0.52)	366 (0.18)	0.307	0.98 (0.82-1.18)
rs783146	CC/CG/GG	571 (0.69)	225 (0.27)	28 (0.03)	1420 (0.71)	528 (0.26)	46 (0.02)	0.265	0.91 (0.76-1.09)
rs2295368	GG/GA/AA	301 (0.37)	402 (0.49)	121 (0.15)	656 (0.33)	1007 (0.50)	333 (0.17)	0.256	1.18 (0.99-1.39)
rs4252135	GG/GT/TT	414 (0.50)	332 (0.40)	77 (0.09)	1008 (0.51)	820 (0.41)	163 (0.08)	0.719	0.99 (0.84-1.16)
rs4252170	TT/TC/CC	683 (0.83)	129 (0.16)	13 (0.02)	1682 (0.84)	298 (0.15)	14 (0.01)	0.789	0.89 (0.72-1.11)
rs783176	AA/AG/GG	556 (0.67)	228 (0.28)	41 (0.05)	1375 (0.69)	557 (0.28)	65 (0.03)	0.065	0.93 (0.79-1.11)

Supplementary Table 5. PLG-coding variants identified in MS patients. Allele frequencies from the NHLBI GO Exome Sequencing Project (ESP), the 1000 Genomes Project (1000G), and The Exome Aggregation Consortium (ExAC) are provided. Chromosomal positions are provided in reference to NCBI Build 37.1. NR, not reported.

Chr	Position	Nucleotide change	Protein change	dbSNP rs/ss ID	Minor allele frequency		
					ESP	1000G	ExAC
6	161127501	A/G	p.K38E	rs73015965	0.004	0.003	0.003
6	161128812	G/A	p.R89K	rs143079629	0.008	0.003	0.007
6	161132146	C/T	p.N110N	rs4757	0.37	0.25	0.26
6	161134069	G/A	p.R153R	rs144153702	0.001	NR	0.001
6	161135876	A/G	p.T200A	rs149145958	0.001	0.001	0.001
6	161137779	T/C	p.C257C	rs14224	0.45	0.45	0.42
6	161137790	G/A	p.R261H	rs4252187	0.003	0.002	0.003
6	161139480	C/T	p.F314F	rs1130656	0.36	0.37	0.38
6	161139857	A/G	p.Q361Q	rs13231	0.26	0.16	0.22
6	161152107	G/A	p.R427R	rs149909079	0.004	0.001	0.003
6	161152155	C/T	p.S443S	ss1467426691	NR	NR	0.00001
6	161152206	T/A	p.S460R	rs116573785	0.01	0.02	0.006
6	161152240	G/A	p.D472N	rs4252125	0.26	0.16	0.22
6	161152257	C/T	p.S477S	rs4699	0.003	0.002	0.001
6	161152807	G/A	p.R490Q	rs140537724	0.002	0.001	0.001
6	161152819	C/T	p.A494V	rs4252128	0.009	0.01	0.008
6	161152837	C/T	p.T500M	rs140970354	0.0004	0.001	0.0002
6	161152858	C/T	p.A507V	rs372603134	0.0002	NR	0.0001
6	161152905	C/T	p.R523W	rs4252129	0.01	0.003	0.007
6	161159619	T/C	p.L618L	rs4252195	0.005	0.001	0.003
6	161162406	T/C	p.A694A	rs4252170	0.06	0.08	0.07
6	161173946	G/T	p.G762G	rs11060	NR	0.31	0.61